CS 134: Networks	Lecture 14, Wednesday, 3/29
Prof. Yaron Singer	Spring 2017

1 Overview

In today's lecture we will discuss influence in the voter model. We will see the connection between this model and the concept of random walks on graphs, and see how to compute influence of a node and select a subset of nodes that maximize influence in this model.

2 Recap: Maximizing Influence

Definition. *For every one of the influence models, we associate a corresponding* **influence function** *which encodes the expected number of nodes influenced by a subset S after t time steps according to the model. That is, if we let $I_t(S)$ be the random variable associated with the number of nodes that changed their opinion after t steps when the initial set of nodes that changed their opinion at time step 0, then the influence function $f^t(S) = \mathbb{E}[I_t(S)]$.*

Influence maximization. Given a graph $G = (V, E)$, an influence function $f : 2^V \to \mathbb{R}_+$ associated with some model and budget $k \in \mathbb{N}_+$, the task of *influence maximization* is to find $S = \text{argmax}_{T:|T|\leq k} f(T)$. In other words, we are finding an initial subset of nodes of size at most k that maximizes the influence in the network.

A useful theorem previously discussed shows that given a monotone submodular influence function f, the so-called "greedy" algorithm will yield a result within $(1 - \frac{1}{e})$ of the optimal solution.

3 The Voter Model

Influence in the Voter Model. In order to formally discuss the idea of influence with regard to the **voter model**, we will use the following terminology. An *opinion* of a node is a binary function with value in $\{0, 1\}$. Initially, a set of nodes S is initialized to 1, and all other nodes have opinion 0. A node is *influenced* at time step t if it had some opinion at time step $t - 1$ and at time step t its opinion is changed (either from 0 to 1 or from 1 to 0). In the models of influence we consider, influence is defined for some finite graph $G = (V, E)$ and time step $t \in \mathbb{N}$.

Note that this is a **non-progressive** model - opinions can flip back and forth between 0 and 1. Previously we have studied **progressive** models, where once a node is "infected" it stays so.

Notation. For each $v \in V$ let $N(v)$ to be the neighbors of v (when we have self-loops, this includes v itself) and $d(v) = |N(v)|$.

Definition. *In the* ***voter model*** *defined on an undirected graph $G = (V, E)$ we assume the graph has self loops* [a]*, and at time t, each node $u \in V$ adopts the opinion its neighbor had at time $t-1$ with probability $1/d(u)$.*

[a] Observe that this is not a structural property of the graph, but rather a more convenient way to describe the model. That is, we can describe an equivalent process on graphs without self loops, but the notation will become messier.

We will be interested in answering the following questions:

- How do we compute the influence of a single node in the network?
- How do we select a set of nodes that are most influential in opinion dynamics networks?

4 Computing Influence in the Voter Model

In order to compute influence in the voter model we will use a few simple facts about random walks in graphs, and then show their connection to the voter model.

4.1 Random Walks

Let us begin by defining a random walk.

Definition. *Let $G = (V, E)$ be a graph with self loops. A random walk of length t that starts at $u \in V$ is a function $r^t : V \to V$ recursively defined as follows: at $t = 0$ we have $r^0(u) = u$ and for any $t \geq 1$: $r^t(u) = v$ with probability $1/d(w)$ for any $v \in N(w)$, where $w = r^{t-1}(u)$.*

A random walk gives us the probability that a node u reaches a node v after t steps when at every step the choice of the next step is performed uniformly at random among the current node's neighbors.

Notation. Given a graph $G = (V, E)$ with self loops, for any $u, v \in V$ we will use $p^t_{u,v}$ to denote the probability that a random walk starting at a node u ends at node v after t steps.

4.2 Matrix Notation

Notation. For convenience we will use matrix notation and assume that the nodes in V are sorted according to some arbitrary order. Throughout the rest of this lecture we will let $n = |V|$ be the number of nodes in the graph. We will associate some unique index $i \in \{1, \ldots, n\}$ with each node v. We will consider $X \subseteq V$. We will overload the notation of X here and say than for an index i associated with v, we will write $i \in X$. We will also say for v associated with index i we have that $v \in X$. So notationally,

$$i \in X \Leftrightarrow v_i \in X.$$

Notation. For some $X \subseteq V$ we use $\mathbb{1}_X$ to denote the vector $(x_1, \ldots, x_n)$ where:

$$x_i = \begin{cases} 1 & i \in X; \\ 0 & \text{otherwise} \end{cases}$$

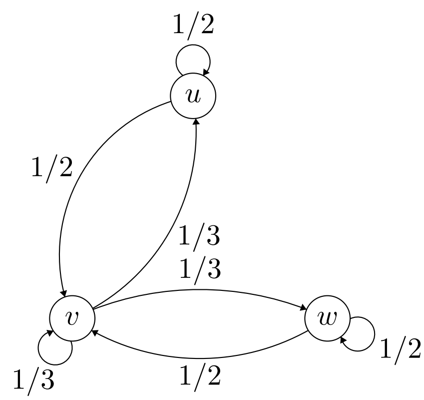
Definition. *For random walks on a given undirected graph $G = (V, E)$, the* **transition matrix** *of the random walk on G is the $n \times n$ matrix M defined as:*

$$M_{u,v} = \begin{cases} 1/d(u) & v \in N(u); \\ 0 & \textit{otherwise} \end{cases}$$

Example: Let $G = (V, E)$ be the undirected graph with nodes $V = \{u, v, w\}$ and edges $E = \{(u, u), (u, v), (v, v), (v, w), (w, w)\}$. Then, the random walk transition matrix is:

$$M = \begin{array}{c c} & \begin{array}{ccc} u & v & w \end{array} \\ \begin{array}{c} u \\ v \\ w \end{array} & \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \end{array}$$

The graph is visualized below:



Proposition 1. *Let $G = (V, E)$ be a graph and M be its transition matrix. Then, for any $u, v \in V$ we have that $p^t_{u,v}$, the probability of going from u to v in t steps, is given by $p^t_{u,v} = \mathbb{1}_u^\intercal M^t \mathbb{1}_v$.*

Proof. The proof is by induction, and is left as an exercise to the reader. $\square$

Example: Consider the graph and transition matrix from the above example.

$$p^1_{u,w} = (1, 0, 0) \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

and:

$$p^1_{w,v} = (0, 0, 1) \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 1/2$$

Lemma 2. *Let $G = (V, E)$ be a graph. For any $u, v \in V$, the probability that u adopts the opinion of v after t steps in the voter model is $p^t_{u,v}$.*

Proof. The proof is by induction on t. For the base case $t = 0$, the claim holds trivially by definition. Assume the claim holds for $t - 1$. For some $t > 0$, we have:

$$
\begin{aligned}
&\mathbb{P}[u \text{ adopts opinion of } v \text{ after } t \text{ steps}] && (1)\\
&= \sum_{w \in N(v)} \mathbb{P}[u \text{ adopts opinion of } w \text{ after } t-1 \text{ steps}]\, \mathbb{P}[w \text{ adopts opinion of } v \text{ after 1 step}] && (2)\\
&= \sum_{w \in N(v)} p_{u,w}^{t-1} p_{w,v}^{1} && (3)\\
&= \sum_{w \in N(v)} \mathbb{1}_u^\intercal M^{t-1} \mathbb{1}_w \mathbb{1}_w^\intercal M \mathbb{1}_v && (4)\\
&= \mathbb{1}_u^\intercal M^t \mathbb{1}_v = p_{u,v}^t && (5)
\end{aligned}
$$

Line 3 follows from the inductive hypothesis, and lines 4 and 5 from proposition 1. □

Theorem 3. *Let $G = (V, E)$. Then for any $S \subseteq V$ we have that $f^t(S) = \mathbb{1}_V^\intercal M^t \mathbb{1}_S$.*

Proof. Notice that when selecting a set of nodes S, the likelihood of a random walk starting at node u ending anywhere in S after t steps is $\sum_{v \in S} p_{u,v}^t$. From Lemma 2 above, this is also the same as the probability that u adopts the opinion of any node in S after t steps. Summing over all nodes $u \in V$ to obtain the total influence of S, i.e. the total influence S has across all nodes in V:

$$f^t(S) \quad = \textstyle\sum_{u \in V} \sum_{v \in S} p_{u,v}^t = \sum_{u \in V} \sum_{v \in S} \mathbb{1}_v^\intercal M^t \mathbb{1}_u = \mathbb{1}_V^\intercal M^t \mathbb{1}_S$$

□

Corollary 4. *Influence in the voter model is an additive function. That is, for every node $v \in V$ there exists a weight $w(v) \in \mathbb{R}_+$ s.t. $f^t(S) = \sum_{v \in S} w(v)$.*

Proof. Consider the double sum formula for $f^t(S)$. We can switch the order of the sums to get $\sum_{v \in S} \sum_{u \in V} p_{u,v}^t = \sum_{v \in S} w(v)$, where $w(v)$ is the total probability across all nodes in V of ending at node v after t steps, or the "weight" of v. This is exactly the result from the corollary above. □

We can now describe a simple greedy algorithm for maximizing influence in the voter model:

ALG 1
input: Graph G, time limit t, budget k
1. Compute the transition matrix M;
2. Sort V s.t. $\mathbb{1}_V^\intercal M^t \mathbb{1}_{v_1} \geq \mathbb{1}_V^\intercal M^t \mathbb{1}_{v_2} \geq \ldots \mathbb{1}_V^\intercal M^t \geq \mathbb{1}_{v_n}$
return: $\{v_1, \ldots, v_k\}$

Theorem 5. *For any graph $G = (V, E)$ and any $t \in \mathbb{N}$ the set returned by* ALG 1 *is the set which maximizes influence in the voter model.*

Proof. Let S be the solution returned by ALG 1. For any set Q of size k, we have that:

$$\mathbb{1}_V^\intercal M^t \mathbb{1}_Q = \sum_{v \in T} \mathbb{1}_V^\intercal M^t \mathbb{1}_v \leq \sum_{v \in S} \mathbb{1}_V^\intercal M^t \mathbb{1}_v$$

where the first equality is due to Theorem 3 and the last inequality is due to the fact that the algorithm takes the nodes with the highest value. $\square$

Thus, it turns out that the "greedy" solution is not just within $1 - \frac{1}{e}$ of the optimal, it **is** optimal, as it is for all additive functions. All results so far were for a general $t \in \mathbb{N}$. An interesting fact about random walks is that after at least n^3 steps the probability of reaching a node u is $(1 - o(1))\left(\frac{d(u)}{2|E|}\right)$ [1]. For $t \geq n^3$ we can therefore use the following algorithm.

ALG 2
input: Graph G, budget k
Sort V s.t. $d(v_1) \geq d(v_2) \geq \ldots \geq d(v_n)$
return: $\{v_1, \ldots, v_k\}$

From the above discussion we have the following theorem:

Theorem 6. *For any graph $G = (V, E)$ and any $t \geq n^3$ the set returned by* ALG 2 *is the set which maximizes influence in the voter model, with high probability.*

4.3 Notes

We saw that in the voter model that maximizing influence reduces to maximizing an *additive* function. A function is additive if $f(S) = \sum_{v \in S} f(v)$. This will be a key property of this model, which often makes it easy to work with.

[1] o(1) is an arbitrarily small value, which approaches 0 as t increases. For $t \geq n^3$, this is small enough to be considered approximately 0